Tissue Bank and Pathology Tools Workspace
Draft Meeting Notes
10/19/04

Meeting Date	**10/19/04 12:00 – 1:00 PM EDT.**
Attendees:	**Facilitator**: *Greg Eley, PhD* **Participants:**

Name	Institution
Ana Holzbach	Mass. General Hospital
Andrew Winter	Northwestern University
Arumani Manisundaram	BAH
Barbara Schaeffer	Dartmouth University
Bob Morell	Wake Forest University
Cathy Counsell	Ardais Corporation
David Aronow	Ardais Corp
Debbie Krupke	The Jackson Laboratory
Derek Walker	Fred Hutchinson CC
Dhugal Bedford	Northwestern University
Emily Burns	BioImagene
Fei Xu	NCI
Greg Eley	BAH
Jack London	Thomas Jefferson Unversity
John Gilbertson	University of Pittsburgh Medical Center
John Speakman	Memorial Sloan Kettering CC
Kyle Allain	Cerner Corporation
Leslie Derr	NCI
Mark Watson	Washington University
Rebecca Crowley	University of Pittsburgh Medical Center
Reechik Chatterjee	BAH
Robert Lanese	Case Western University
Steve O'Krepky	Rose Li Associates
Susanne Ragg	University of Indiana
Tara McSherry	University of Pennsylvania

Agenda	1. Welcome, Logistical Issues – Greg Eley - BAH 2. caBIG Compatibility Document Overview – Arumani Manisundaram - BAH
1. Welcome, Logistical Issues	After welcoming everyone to the teleconference, Greg Eley took attendance and then turned the meeting over to Arumani Manisundaram (workspace lead for the Architecture Workspace), who began a discussion of the caBIG Compatibility document. Members of the TBPT workspace have reviewed this document over the last couple of weeks.

Tissue Bank and Pathology Tools Workspace
Draft Meeting Notes
10/19/04

2. caBIG Compatibility Document Overview

The caBIG compatibility document provides a description of caBIG compatibility, a description of the various levels of caBIG compatibility, and provides examples of each level of compatibility. Arumani began his overview of the guidelines document by the caBIG Compatibility Guidelines to the caBIG Principles. The caBIG principles describe the sociologic aspects of the processes and artifacts developed with NCI/caBIG resources. The caBIG Principles are four fold:

- Open Source
- Open Development
- Open Access
- Federated

Whereas, the caBIG Guidelines describe an Architecture and Data Representation that facilitate semantic and syntactic interoperability between information systems. Thus, a cancer center employing a legacy solution or a vendor developing to the caBIG Guidelines can still be compliant without adhering the caBIG Principles. Currently, the caBIG Compatibility Guidelines document is a living document and is publicly available for consumption or comment. Arumani stressed that the NCI, Architecture, and Vocabularies and Common Data Elements Workspaces would like to receive feedback on the document.

To ensure that the compatibility guidelines are appropriate, the Architecture Workspace has been gathering use cases from the caBIG community. These use cases have helped the cross-cutting domains to begin to understand the challenges that need to be overcome in order to achieve a functioning grid.

Many questions were raised regarding the caBIG compatibility document. The types of questions can be broken down into 2 major categories: 1) application requirements, 2) data requirements

APPLICATION REQUIREMENTS:

- How is a tool, grid-enabled?
    - The Architecture group will specify the grid technologies which will be based on the Globus and OGSA-DAI Toolkits. The application must be able to leverage CDEs and standard terminologies for the purpose of sharing information across the grid.
- Will the Architecture Workspace provide a Grid Toolkit to facilitate grid enabling of applications?
    - At the very least, the Architecture workspace will provide the specifications for the caBIG. It is possible that the Architecture Workspace will develop a toolkit to act as the middle layer between an application and the grid

Tissue Bank and Pathology Tools Workspace
Draft Meeting Notes
10/19/04

technologies.

- Will a Gold Level application be required to handle data at all levels of compatibility, or data compliant only at the gold level?
    - The answer to that is unclear. The current understanding of Data compliance vs. Application compliance is not well defined.
- To reach caBIG compliance, will all tools need to be completely modeled in UML? Will other models suffice? Are models even required?
    - The only modeling specified in the compatibility guidelines are for Information Models. The models should be in a common modeling language and represent the Object Models and how the data interacts. Common models include: Class diagrams, Sequence diagrams, and Collaboration diagrams.

DATA REQUIREMENTS:

- What is the difference between silver and gold level compatibility for data?
    - (The answer to this questions needs to be defined. There was much debate over this question)
    - Silver: use CDEs to represent data, use commonly agreed to messaging protocols
    - Gold: Silver + associate meta-data with the actual data + share data over the grid
- It will be extremely difficult to expect all participating Centers to collect and share the same information. How do we allow for slight variations in CDEs when sharing data on the grid?
    - (The answer to this questions needs to be defined. There was much debate over this question)
    - One option is to agree to a CDE with a variety of possible data types and values. Then each Center can represent their data to the extent possible, given their institutional and process constraints. There may be missing data, but all available data will be represented in a common way.

Regarding the technology for the Grid, it was stated that a well-described API provides an interoperable component. Arumani stated that vocabulary needs to be based on a common messaging system.

CDEs should be harmonized across domain workspaces. To achieve gold compatibility, an application has to be completely interoperable with the data grid. To move from silver to gold compatibility, one does not have to re-engineer their application. Gold compatibility is just a logical progression from silver. Gold will allow for more complete search capabilities.

Tissue Bank and Pathology Tools Workspace
Draft Meeting Notes
10/19/04

It was mentioned by one of the cancer centers that there are really two grids. One is the caBIG data grid; the second could be for the metadata. caDSR serves the purpose of creating unique identifiers for data objects. The metadata is tracked by using the unique identifier. The "wrapper" on top of the data sources that maps to the grid is a characteristic of gold compatibility. The tissue bank tool being grid enabled pulls data from the back end. The instructions to do this can be provided from the API model.

It was also stated in the meeting that some type of clarification is needed in terms of examples to clearly show what is gold and what is silver compatibility for both applications and data.

Additional discussions regarding the data representation and silver/gold compatibility should be addressed in the TBPT VCDE SIG. The VCDE SIG should provide a structured list of requirements/constraints/clarifications for the VCDE and Architecture Workspaces. This will provide a vehicle for adding value to the grid and ensuring compatibility.

The meeting ended with Arumani providing closing remarks and encouraging all participants to work together to define their needs, and to provide feedback to the Architecture and VCDE Workspaces on caBIG compatibility.

The next meeting will be held on 2 November, at 1200 ET. The topic of the teleconference will be caTISSUE.

Action Items:

Individual Responsible	Action Item	Due Date	Notes

Tissue Bank and Pathology Tools Workspace
Draft Meeting Notes
10/19/04
